1 (2000-R)

PILOT ASSESSMENT ASPECTS OF SIMULATION

By George E. Cooper and Fred J. Drinkwater III

Ames Research Center, NASA
Moffett Field, California, U. S. A. 94035

INTRODUCTION

This lead paper on the pilot assessment aspects of flight simulation discusses in greater detail some of the problems introduced in AGARD Report 567, "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities" (1). The important function of a lead paper on pilot assessment is to introduce the critical questions raised by pilots so they can be examined and discussed with the aim of developing solutions and improved understanding. Some answers are proposed that may in themselves be controversial and stimulate further discussion.

Our major difficulty in the application and utilization of pilot assessment in simulation is that there are no simple black and white answers for many of our problems, and continuing communication between pilots and engineers is essential. This is not so difficult to understand when we realize that with simulation we are seeking answers with only part of the tools for the job, and we are using the human element, namely the pilot, to bridge the gap between simulation on one hand, and the final flight application on the other. It is important, therefore, that we involve the pilot as early as possible in developing a piloted simulation program.

In order to bring out the pilot's viewpoint, we have reviewed common pilot gripes and complaints arising from simulation experiences and selected a number of questions or problems that we believe focus discussion on areas of maximum interest and concern. We consider first the apparent primary concern of pilots participating in simulation work, and next the questions related to the pilots' actual participation in the planning and conduct of experiments, the simulation situation in terms of the facility being used, and the analysis and reporting of results.

1. THE PILOT'S PRIMARY CONCERN

The pilot involved in a simulation experiment is concerned primarily with the adequacy or fidelity of the simulation with respect to the objectives of the program. Stated another way, this factor is the degree to which the pilot is expected to extrapolate simulation results to the actual flight situation.

In previous AGARD reports or papers (1 and 2), care has been taken not to relate the usefulness of simulation experiments with simulator sophistication. It has been shown that the rudimentary simulator, if properly used, can solve many of our problems or provide the needed guidance for reaching their solution. It is only natural when given a useful tool that we try to extend its capability to the maximum. By obtaining increasingly definitive answers we can take much of the risk out of new vehicle development. In so doing, however, we are apt to press the pilot into providing ratings and evaluations that are beyond his ability to extrapolate. At times, this may tend to shake his confidence in the whole simulation process. The planning and conduct of simulator experiments to obtain the full confidence of the evaluating pilots is discussed in reference 1. Other aspects of pilot participation discussed here relate to program definition, the pilot as a subject, the pilot as an evaluator, simulator validation, and pilot adaptation.

2. PILOT PARTICIPATION

2.1 Program Definition

We must first determine the program objective. What is it we wish to learn? What problem is to be solved? Are we seeking approximations to learn more about the relative effects of specific parameters or interactions? Or are we concerned about handling qualities in a specific critical task? What are the important aspects of the program? It may not be essential that a pilot be directly involved in program definition? However, his participation is desirable, for many simulation objectives will have been identified by pilot observation or experience. Pilots should be concerned with task definition, particularly with determination of the critical task and the environmental conditions associated with it.

Definition of the pilot's role in the simulation experiment also is important, for it is a basis for determining the extent of pilot participation in the rest of the program, i.e., planning, the execution of the experiment, and the analysis and presentation of results. Some hold that pilot extrapolation is the most important benefit to be obtained from the simulation experiment and that its value lies in the fact that the pilots' experience enhances his ability to interpret limited information in terms of a real aircraft operating in a realistic task as part of a defined mission. The other view states that pilot extrapolation is to be avoided at all costs if we wish to achieve satisfactory data. As pilots, we consider that these statements are directed at different uses for the simulator and therefore are not necessarily in opposition. If the simulation is primarily to obtain engineering or test pilot observations or assessment of the suitability of the simulated airplane characteristics, then it should be considered as a "pilot's tool" enabling him to perform maneuvers and tasks that flight test experience has shown will permit appropriate observations and assessment. His extrapolation to the real world, his opinion of the fidelity of the simulation, and the value of the results based on flight experience are most important. On the other hand, the simulator becomes the "experimenter's tool" when he is interested in measuring the pilot's performance and observing the pilot's reaction, as a subject, to operational situations, vehicle characteristics, handling qualities, or other factors. In this case, the pilot is usually one of a number of subjects and his observations are less important than the measured test parameters. This distinction in the use of research simulators is the first step in establishing the degree to which pilot extrapolation is involved.

2.2 The Pilot as a Subject

What are the important factors when the pilot is to be used as a subject or a task performer rather than an evaluator? Is it essential that the pilot use the same techniques as he would in the aircraft? Is task performance the primary result to be sought? If pilot performance is the primary data output, how is pilot input workload taken into account? This last question identifies the pilot's concern about the use of measured performance alone as the criterion for simulation results. Do experiments in which the pilot is a subject require the highest degree of simulator and task fidelity for results to be applicable to the flight situation? Is not this precision and completeness exactly what is required in the ultimate training simulator? The pilot must have the greatest fidelity and vehicle duplication with realistic task and workload demands if he is to be evaluated purely on the basis of his performance.

What about the human factors researcher? Does he require the ultimate in simulation sophistication in order to make any contribution? We propose that the sophisticated training simulator does offer a most valuable tool for human factors studies and should be used to a much greater degree in the study of pilot performance. On the other hand, what about the human performance experimenter who does not have access to such sophisticated simulation facilities? Our answer is that basic research in *human* performance capabilities is essential to understanding the human problem, but that the extrapolation by the scientist of isolated operator performance and subject reaction experiments to the aeronautical task is subject to severe limitations.

How then can simple compensatory tasks using only the basic elements of the controller and controlled elements be presented and interpreted? If the answer is to define the pilot transfer function, then one must know how to apply the pilot model effectively. Useful areas of application can be visualized but, to the pilot, the introduction of complex interactions in real life without adequately defining pilot workload restricts the application of the results. Unless the full task is simulated with very high fidelity, the only practical way to relate pilot workload to output performance is by pilot assessment. An intriguing question continues to arise, however. Is not there a common denominator in pilot workload that would permit substitution of an artificial secondary task for the actual task during the simulation until a full task (sophisticated training) simulator becomes available? The pilot's answer would probably be that if there is a way to create full task workload, it would be related to the time the pilot is distracted from the primary control task. Pilot guidance and suggestions are needed to facilitate more profitable work in this area.

2.3 The Pilot as an Evaluator

With the concept of the simulator as the "pilot's tool," the pilot's role will emphasize assessment. If the real value of simulation is the ability to extrapolate to the actual flight situation, then who will make the extrapolation, the experimenter or the pilot? The engineer may devise a series of part task simulations to obtain results as separate task ratings that he can then attempt to extrapolate to the complete task (or flight phase), or he may ask the pilot to provide the extrapolation by assigning either a flight phase or a composite rating based on one or a series of part task experiences.

As an example, assume that we have a simulation that allows the pilots to perform any stated maneuver (banking, turning, pitching, turn reversals, etc.) but does not have the capability for extended tracking of a maneuvering target. Assume further that a simulation objective is a definition of stability and control characteristics that will establish criteria for an air superiority fighter. The experimenter could ask the pilot to perform each specified maneuver in turn and then take the separate ratings given for each maneuver and attempt to integrate into a common rating for the characteristics of an air superiority fighter. However, this approach does not take advantage of all the tools available: Why not give the pilot freedom to integrate these various simple maneuvers from his experience and give him the task of equating to the real world situation. Pilot extrapolation usually is to be favored over extrapolations by the experimenter unless the objectives can be limited to relative effects of parametric changes or the experimenter is provided with guidance for more definitive tests to follow.

In attempting to identify the extent of pilot extrapolation in a given program, we will probably find that it is not possible to define explicitly what the limitations are or what agreements must be reached; instead, we must settle for the general statement that the application of results from simulator programs should not exceed the confidence with which the pilot accepts the limitations in the simulation and his own extrapolation factor. There may be a relationship between the extent of the required pilot extrapolation and the manner in which data are extracted from him. It is important to recognize the different concepts represented by task, composite, and flight-phase ratings that imply differing degrees of extrapolation by a pilot. In providing a task rating, the pilot is assessing only the task he is actually asked to perform. Flight-phase rating, on the other hand, implies that not all tasks and flight conditions applicable to the flight phase have been provided for pilot assessment. This requires the pilot to interpret and assess the effect of what is missing. In this case, pilot extrapolation is maximum. An intermediate level of extrapolation is introduced by a composite rating for which the pilot first rates each task separately and then assigns an overall rating for the flight phase.

2.4 Simulator Validation

The completeness of any simulation and the degree of fidelity with which a given vehicle, environment, and task are reproduced is important to the pilot, but he may not become involved with them until he is first required to operate the simulation. It is usually easier for the pilot to determine gross errors or significant items or parameters that are missing than it is for him to determine if the simulation is explicitly correct. He can detect very quickly conflicting information, erroneous time constants, and excessive lag between visual and motion cues or between external visual displays and instrument readings when they are associated with a particular pilot control input. In many cases, he may be able to offer advice as to the relative importance of various parameters in maintaining acceptable fidelity of a given simulation program. Most pilots feel that fidelity to the equations of motion is the responsibility of the experimenter or simulation engineer.

The question of simulation fidelity and checkout leads to one of the more vociferous pilot complaints: being asked to begin a simulation data-gathering phase with a simulator that is not completely validated. It is not uncommon to encounter tenfold errors in scaling, instruments hooked up backwards, parameters not introduced properly, and other problems associated with human error. We may accept it as part of the simulation pilot's job to uncover such errors and to assist in the validation of any simulation, but these chores should certainly precede the initiation of a data-gathering phase.

What the pilot asks, however, is that as many as possible of the validation exercises be performed before he enters the picture. This applies not only to ensuring anticipated response to control inputs through proper reaction of the important instruments, but to verification of the scaling, resolution, contrast, etc., of visual systems, and the measurement of actual accelerations occurring in the cockpit of motion simulators.

2.5 Pilot Adaptation

The use of a reference airplane with which the pilot is familiar can greatly assist him in adapting to a particular simulation, particularly if the simulation is unsophisticated and incomplete in certain respects. Knowledge of his reference certainly will facilitate the process of extrapolation, and performance and workload with a reference airplane will provide data for the pilot in evaluating his performance and workload with the vehicle under study. The length of time required to adapt to a simulator (i.e., to provide consistent performance and to apply the same piloting techniques as he would in flight) might be assumed to be inversely proportional to the degree of sophistication involved in the simulation. There is a relationship, but it is not quite this simple. In the first place, if the pilot is in a familiar environment with respect to the cockpit interface (instruments, controls, displays, etc.) he can begin using them efficiently much more quickly. If, however, important cues or bits of information are lacking or misplaced, it may take him considerable time to adjust to this situation and to learn to use alternate or substitute cues in the performance of the task. For example, in the authors' experience, during some early simulation studies involving an outside world visual display scene, the resolution in the TV picture and height perception were not sufficient for the pilot to obtain all his normal VFR cues in the same manner as he did in VFR flight. As a result, he was attempting to evaluate landing performance in normal VFR under conditions approximating poor visibility. This situation necessitated considerable practice and attention to variation in aircraft attitude and indistinct runway markings to obtain even a fraction of the information required to land. Sufficient trial-and-error runs had to be performed to develop substitute means for judging the flare height simply because the height cues were not as clear cut as those encountered in actual flight. These limitations might be considered by some as invalidating the use of such a simulator; however, when the pilots were given the opportunity to go through the lengthy familiarization process, they became confident in use of the simulator and felt they were ultimately using techniques that could be related directly to those used in flight. Can we develop rules or guidelines to assist in determining how much pilot familiarization is necessary? Or should the pilot determine when he is ready to provide data runs? Can this decision be made arbitrarily on a common basis for all participating pilots? The attainment of a performance plateau is a useful criterion but in using it, the experimenter must be sure that the pilot is striving for maximum performance during his familiarization runs. Quite often, it should be noted, this is not the pilot's objective during familiarization; he may employ a series of different maneuvers and tasks, "challenging" the simulation and configuration to determine what its characteristics really are.

3. SIMULATION SITUATION

In general, the equipment available imposes certain constraints on the simulation situation. The simulator used during the development of a new aircraft may evolve from a rudimentary prototype of a thoroughly sophisticated and advanced piece of equipment. At each stage, it is important that the simulator be used in a manner compatible with its degree of sophistication and fidelity. Typical questions that can be posed in this connection are: Can compromises in the fidelity of cockpit presentations be compensated for by additional simulator training and preparation by the pilot? Are there any general rules that pilots can provide relative to the degree of fidelity required in the cockpit interface features in a given simulation situation?

3.1 Cockpit Interface Features

Over the years, pilots have complained about unfamiliar instrument arrangements, scaling, or substitute display features. Attempts have sometimes been made to "make do" with control systems having excessive friction, high break-out forces, improper centering, or other characteristics that interfere to a large extent with the pilot's ability to relate the simulation to an aircraft. Such distracting features must impair, to some degree, the pilot's ability to focus properly on the variables in question. For example, many pilots would find the illusion of simulating a high performance fighter completely destroyed when faced with a wheel-type control and high force gradients. The measurements of aircraft stability and control can be misleading unless the influence of control system friction is known and appreciated. A pilot's assessment, on the other hand, would normally be based on the "apparent stability" as influenced by control system characteristics. The pilot should not be expected to separate the effects of excessive friction from the true aerodynamic characteristic. If anything, pilots tend to be too adaptable and cooperative in such cases. Too often the situation is recognized only after the program is fully developed and ready for initial pilot evaluation. Many other controversial questions could be asked concerning simulator control system and selector characteristics. The following simple rules may be helpful in the definition of these cockpit interface features.

1. Whenever possible, provide a cockpit design with which the pilot is reasonably familiar, unless the design features are to be variables in the study.

2. Minimize unrealistic control characteristics such as poor centering, high friction, and poor readability of instruments before starting an experiment.

3. Allow the pilot to participate in the cockpit layout, to evaluate friction, lag, and breakout, and to select control gearing and/or sensitivities whenever he feels they will seriously affect his performance.

3.2 External Visual Display

In recent years, the external visual scene has become increasingly important in the use of flight simulation. When is the external visual scene necessary, and what is its value to pilot assessment in simulation? What are the sources of complaints or limitations concerning current external scene visual displays?

At the very least, we can say that the external visual scene is required whenever external information is required to complete the defined task or flight phase. For some programs, the visual display may merely provide the pilot with increased realism and a greater capacity for evaluating his own performance. The external visual scene is required when it forms the basis of the most critical task to be assessed by the pilot. Thus, it may be more important to define critical tasks involving external vision. Two situations in which the external visual scene may become critical are in the terminal area VFR operation and the VFR landing. Increasing air traffic and reliance on visual contact to avoid collision require the pilot to fly a precise path on instruments while maintaining a continuous lookout for other aircraft. Either additional crewmembers must be available to watch for other traffic or a distractive influence must be introduced to represent the additional tasks. Other situations in which the external visual scene is necessary are air-to-air tracking or other visual positioning tasks such as formation flying or air-to-air refueling.

What are the primary requirements for improving external visual displays? The visual landing task and the associated low-altitude transition from instruments to visual control have caused pilots to ask for greater and greater improvements in height perception from the displays. For this task, pilots are more than willing to accept a reduction in the available VFR ceiling capability if they can achieve improved resolution. Pilot guidance is useful in determining the important factors in compromises of this nature. Another example is the requirement for 360° flyaround capability. This has been primarily a training requirement and its importance to research has probably been overrated. However, it would be welcome on a research simulator if it did not introduce a more serious compromise in another area. A wide-angle projection capability would certainly be desirable, but it has been of less importance than resolution for conventional approach and landing work. Peripheral vision cues normally observed through side windows become more important to the pilot when forward vision cues become reduced or more difficult to interpret. For V/STOL work, however, where very high turn rates are involved as a consequence of the slow speeds, an enlarged visual scene appears to become much more important and could very well be the limiting factor in this type of simulation.

Are there advantages in color over black and white displays? We have been extremely pleased with our Redifon color projection system, the acceptability is believed to result more from the degree of resolution and depth perception achieved than from color per se. There are common approach and landing scenes in actual operation that are almost devoid of color. The runway itself, from which the pilot gets the majority of height, flare, and touchdown cues, is in black, white, or gray. Color in the surrounding terrain and definition of the approach and runway lighting with color coding, for example, are important to the pilot when they provide usable cues. The high-resolution information needed by the pilot is primarily in black and white unless, of course, a VASI system is used. The attainment of high resolution must logically start with the visual display model being used, and the larger the model the greater detail that can be put into it. Then a high-resolution TV camera and projection system do their part.

3.3 Simulator Motion

In some cases, simulator motion has meant moving the cockpit about to simulate the general feel of aircraft movement through the vibration, shaking, and small accelerations to which the pilot is accustomed, and it generally relates to actual flight experience in the form of minor to severe disturbances. In other cases, it has meant more accurate reproduction of actual accelerations at the pilot's station under given environmental conditions or in response to specific controlled maneuvers to provide the important cues. To obtain the latter conditions we have progressed from single-degree-of-freedom to six-degrees-of-freedom motion on our simulators while only occasionally achieving the realism the pilot looks for. We often are not selective enough in determining which motions are important to the study and which can be simulated most accurately without introducing disturbing and unrealistic motions. The second problem, of course, is that sustained acceleration is not possible without a very large translation capability. A pilot simply cannot accept erroneous acceleration cues that are much above his perception level. Whenever maneuvers requiring washout or erroneous cues are necessitated by the selected task, care must be taken to keep erroneous accelerations low. There is no substitute for high fidelity when motion is used to simulate important *cues* because a lack of fidelity changes them to *distractions*. Some loss in fidelity, although undesirable, can be accepted when distractive motions are introduced.

A way to achieve the greatest realism for small movement and effort is by reproducing the relatively high-frequency accelerations that occur in turbulence. Because these are basically disturbances rather than cues, they are more easily simulated by limited-motion simulators. Longer term drafts and wind shear effects are important to the fidelity of the simulation even though not reproduced in motion. If we are forced to be selective in our motion simulation to minimize erroneous accelerations on the pilot, the motion considered next most easily attainable and effective in improving realism is the normal acceleration in response to pitch control. Again, such accelerations are not possible as sustained values, but normal acceleration is the most active acceleration encountered by a pilot during normal manual flight operations. More or less continuous maneuvering in pitch control is performed by a pilot and results in a continuing feedback of normal acceleration. It has been noted that this acceleration feedback mode provides a significant function in reducing pilot adaptation time on a simulator. The tendency to overcontrol is greatly reduced. Lateral-directional motions are the most difficult to achieve with realism while at the same time minimizing the effects of erroneous feedback from washout. The major success achieved in this type of motion is attributed to the provision of large lateral translation that enables pure side acceleration without the interrelated disorienting washout associated with banking to achieve side force. The best solution in many cases lies in restricting simulator bank angles and rates to reduce erroneous side accelerations without seriously impairing the normal banking cue. The banking (rolling) cue is probably overemphasized and overrated in many simulations.

In general, motion inputs for cues must be correct in timing or phasing or should be left out of the simulation. Attenuation in amplitude of acceleration may often be required and can be acceptable, providing the rate of onset is distinguishable and appears correct. Lags and delays in response can be intolerable, particularly when they can be correlated with corresponding instrument indications or external visual cues.

Many pilot complaints could be overcome if a few simple rules could be observed:

1. Determine the motions considered most likely to influence the simulator realism and pilot performance.
2. Maximize the important cues to the extent possible while minimizing any erroneous accelerations, even if it means not using all the degrees of motion freedom available.
3. Provide simulated turbulence based on actual turbulence profiles or at least proper distribution of energy.

3.4 Simulated Environment

When do environmental disturbances become important or necessary to a simulation program? What are the pilot's chief concerns relative to simulation of crosswind, wind shear, and turbulence? Environmental disturbances can contribute to, or define, the most critical task with which the pilot has to cope; or they may establish the operational limitations beyond which safe flight cannot be assured. For these reasons, fidelity in the aerodynamics of the simulation of these effects is important even though all the motions and accelerations cannot be achieved.

3.5 Pilot Stress

Can a simulator pilot adequately evaluate the real world surprise or distraction associated with transient disturbances? Or to what extent should surprise and unanticipated disturbances be provided in the simulated task? Actually there are many deficiencies in the way the surprise element has been introduced in the past during handling-qualities evaluations, certification testing, and pilot training.

Accelerated service testing of aircraft provides a high utilization period during which system failures and unanticipated problems are "encouraged" to occur, primarily so that material and mechanical deficiencies can be anticipated and corrected before normal operation is begun. Perhaps we ultimately should use the sophisticated training simulator to provide a service test feature for the pilot and machine by studying their response to truly unanticipated failures and disturbances under a variety of operational situations.

4. ANALYSIS AND REPORTING

4.1 Pilot Assessment Data

Aside from the part a pilot may play in planning, establishing, and operating a simulation experiment, his primary contribution usually consists of performance outputs and assessment data. We can point out two extremes in simulation programs: the use of only the pilot performance data, and the use of only the pilot assessment in terms of a rating and comment. Test pilots are usually more critical of how their assessment data are used than their performance data. The use of performance results from part task or rudimentary simulations, however, can be accepted by pilots only if results are extrapolated in context with the limitations in the simulation. How can we ensure that this is done?

Let us first turn our attention directly to four ways of presenting the pilot assessment data and examine several controversial aspects. What are the chief concerns on the part of the pilots relative to the presentation and analysis of simulation data? What arises when a pilot is not allowed to participate in the analysis or does not have the opportunity to do so? The latter situation can arise for a number of reasons, including those beyond the control of the experimenter such as pilot unavailability. If it is not physically possible for the pilot to participate in analysis of results, this fact should be anticipated by a commitment to use pilot comments or to request a summary statement by the participating pilot.

One method of presenting assessment data is to give pilot ratings for a specific handling-quality parameter and flight phase. This method provides guidance and insight into relative effects of changes in specific parameters, but it does not always address the actual level of handling qualities; assessment data are incomplete without supplementary comment.

Second, handling-quality task ratings can be presented with comment data. This method usually results in the most objective treatment of the assessment data and enables the reader to analyze pilot assessment results himself to verify the conclusions reached. This procedure has the additional advantage of influencing the care with which pilots record their ratings and comments.

A third procedure is to provide average or mean pilot ratings, rather than individual data, and selected pilot comments. The result is a summary of pilot assessment data as compiled by the experimenter or the project pilot. This summary reduces the bulk of pilot data and, with careful editing, constitutes a simplified logical analysis of pilot conclusions. It has some advantage in simplification and clarity, but the reader does not have the freedom, if he so desires, to examine all the raw data. As with any experimental process, this method's acceptability probably depends on the intended use of the data or the results of the experiment. When results and comments of different pilots are similar, some editing and consolidation may be acceptable. One problem we see here is that trends that may be evident from analysis of individual pilot results will not be apparent in the data summary. This method may also obscure trends that may not be significant for the basic study objective but could be of interest in other applications of the data. On the other hand, during many "short-look" programs, it is not possible for a single pilot to observe and comment on all significant factors that may contribute to the results and conclusions.

The fourth method of presenting pilot assessment results is by setting aside a pilot discussion section that summarizes pilot observations, comments, analysis, discussion, and conclusions. This method is most often used when one or more pilots share in the authorship of a report, and has the advantage of summarizing assessment data that might otherwise be considered too detailed. In this case, responsibility for editing and objectivity rests with the pilot.

Regardless of the concerns expressed, limitations inherent in simulation, and the variety of ways in which pilot's assessment can be used, more emphasis should be given to the trained pilot's capability as an observer and reporter.

5. SUMMARY COMMENTS AND QUESTIONS

1. The primary factor concerning the pilot who must provide assessment data in a simulation experiment may be stated as: the adequacy or fidelity of the simulation with respect to the objectives of the program; that is, the amount of extrapolation expected of him.
2. Usefulness of a simulator is not necessarily related to its sophistication.
3. The distinction between simulator use as an "experimenter's tool" or as a "pilot's tool" is suggested for clarifying the pilot's role and participation.
4. Is the concept of *task, flight-phase,* and *composite* ratings helpful in identifying the extent of extrapolation required?
5. Are there useful suggestions pilots can offer in the area of pilot workload and output performance measurements?
6. What is or should be the pilot's role in simulator validation?
7. What factors influence pilot adaptation to a simulator?
8. At risk of oversimplification, several rules and pilot observations are given with respect to the interface features of the control systems, flight instrumentation, and the use of cockpit motion and outside visual displays.
9. Increased research application of the sophisticated training simulator is expected. One potential new role is suggested in the area of operational research involving the use of true stress and reaction factors associated with unanticipated failures and disturbances – in effect, accelerated service test of the man and machine.
10. The trained pilot's capability as observer and reporter should be used to fullest advantage.

REFERENCES

1. Cooper, George E.; and Harper, Robert P., Jr.: The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities. AGARD Rep. 567, April 1969.
2. Cooper, George E.: The Use of Piloted Flight Simulators in Take-Off and Landing Research. Paper presented to the Take-Off and Landing Specialists Meeting Flight Mechanics Panel of Advisory Group for Aeronautical Research and Development, Paris, France, January 14–18, 1963.

LEAD DISCUSSION

by

J. Pinet
S.N.I.A.S., France

After reading the "Lead Paper" prepared by Messrs. Cooper and Drinkwater, I was most embarrassed trying to find matter for discussion. All the questions raised were those we already had asked ourselves during Concorde or Airbus Simulation studies, and we too would like to hear replies from the people "in the know," for they would help us very much. In fact, as a discusser on a subject with which I agree, I will endeavor to answer the questions raised in this paper to the extent where we have some experience on the subject, wishing to avoid any "theoretical" discussion.

Let us take the various sections in succession.

The Pilot's Primary Concern

Just a comment - I think that a simulator should be more and more sophisticated as progress is made in the definition of the thing to be simulated. Thus, the study simulator for one aircraft type can be less sophisticated than that for a given aircraft.

Inversely, the more succinct the simulation, the greater should be the qualifications of the pilots assigned to the simulation tests, in terms of their experience on the involved aircraft types and related simulation.

PILOT PARTICIPATION

Program Definition

At S.N.I.A.S., we have selected a solution allowing an excellent efficiency in simulation testing. The Engineering and Flight Test Departments are jointly in charge of defining the standards of simulation necessary for the studies as well as the programs. More particularly, these involve an engineer from Engineering and a Flight Test pilot assigned to the simulation. Then, the task of performing the programs is given to the pilots or crews who have been selected, as far as practicable, in the light of their experience and capacity to evaluate and judge the tested cases, although their availability at the time is the fundamental criterion for their selection!

The Pilot as a Subject

Generally speaking, we have no programs devoted to pure research, and most of the time the objective is to define the best characteristics for a given system, to evaluate the benefits to be gained from a new system, or to evaluate the acceptability of a new configuration. The methods used are based on what we call "common sense". In the absence of positive criteria, we try to place the pilot, or the crew, in a situation as close as possible to reality. We also select the crew to match the program requirements, giving great consideration to their opinions which we have, let us say, "calibrated" with experience, while also checking their results with conventional performance recordings.

To date, this mostly empirical method has been satisfactory for us, but I do imagine that it cannot always apply to fundamental research.

Further, I wish to say that we have been favored by our simulator in Toulouse, which has a broad range of possibilities. However, we have experienced difficulties in establishing ratings, and we have been led to define a rating system for our own use, taking implicitly into account the workload. This system is described in the Appendix. It has no other pretention but to be useful for our tests. Roughly, it takes into account the pilot's "inputs" (required attention) and his "outputs" (required skill). Maybe this rating system will give you ideas on a better system, and I would be very happy to hear your criticisms and comments on this subject.

The Pilot as an Evaluator

I am as much in favor of partial simulations on a well defined task to:
- determine the "basic" pilot behavior
- compare several definitions of systems to be tested
- or simply try to "clarify the situation"

as I do not trust results from partial simulations to make an overall judgment. It is the work of a team of engineers and pilots which allows us to make a synthesis of the results obtained, and this method is the only one which seems sure today. Once more, I emphasize the necessity of a pilot-engineer "basic team".

Simulator Validation

The problem of simulator validation is relatively simple when the simulated airplane is already flying. We have developed a method which seems efficient:
- static and dynamic comparison of the aircraft and simulator transfer

functions, and readjustment of the aerodynamic coefficients by means of a digital computer program
- readjustment of the simulator settings
- check of the simulator by the pilots using characteristic maneuvers.

This method concerns the flight mechanics of the simulation and the pilot is in the loop at the very beginning of the checks after the aerodynamic coefficients have been reset.

Simulator validation is very important for motion cues. We had great difficulties in adjusting our cabin motion system, and my paper on "Cockpit Environment" comments on some of them. The fact that we had a pilot assigned to the simulation and judging the simulator at the beginning of each program was most helpful.

But even before Concorde had flown, when we were trying to validate the simulator with the characteristics of another aircraft we knew very well, because of the tremendous amount of numerical programming for a given aircraft and the necessity for rapid progress in our tests, we were led to give up this method of attempting to validate the simulator.

On the subject of simulator validation, once more I am in favor of a joint assessment from pilots and engineers. I can give you an example where the pilots gave a wrong opinion of the simulation. Before Concorde flew, a kind of "stick-boost" was introduced in the roll control system. On the simulator an abrupt roll motion was felt every time the pilot introduced a roll command. Since this motion was felt to be unrealistic, a filter was introduced in the roll axis of the simulator's motion system. When the aircraft flew, this phenomenon was at once felt and we removed the filter - and the "stick-boost."

Pilot Adaptation

It should be noted that the introduction of new pilots into our simulator activities is always done under the supervision of an experienced pilot who corrects and judges the newcomers behavior.

Some pilots can be asked to develop evaluation data at the same time they are conducting runs for performance measurements. But these pilots must be chosen carefully according to their experience and their ability to conduct such tests.

SIMULATION SITUATION

I think that the essential parts of what I could say on this subject are included in my paper on "Cockpit Environment."

ANALYSIS AND REPORTING

I am in favor of a combination of the second method (handling quality task ratings with comment data) and the third method (mean ratings and selected comments) according to the tests carried on and to the available pilots selected for the assessments. This gives good results when the pilots are "calibrated."

SUMMARY COMMENTS AND QUESTIONS

In the Appendix of this Lead Discussion and in my paper on "Cockpit Environment" I suggest, in response to the questions raised by Messrs. Cooper and Drinkwater, several answers which apply in a given type of simulation - an aircraft type simulation.

I want to emphasize the necessity of having a kind of nucleus of pilots and engineers who work very closely together on the simulation programs. These very smart people must have a lot of "common sense," before the time when they will have the opportunity and the chance to use a mathematical pilot.

APPENDIX - HANDLING QUALITIES RATINGS

1. PURPOSE

Up to now, it has not been possible to find sufficiently proven and reliable criteria to define the quality of an aircraft's behavior associated with its pilot. Some work has been done worldwide trying to define such criteria. At Sud-Aviation, time, people and money have been lacking to successfully carry out such long and delicate studies. We had to satisfy ourselves with what is universally used with more or less success.

If this rating were actually representative of the aircraft's behavior, it would allow for quick classification of the configurations, much quicker than through processing and interpreting of the detailed pilots' opinions; for, in this latter instance, if the tests and configurations are numerous, the volume of work is tremendous and incompatible with our present possibilities.

Thus the rating offers, of course, a big advantage for the simulator studies by allowing, whenever possible, for quick processing of the numerous cases being tested.

2. PRESENT RATING SCALES

There are three main pilot rating scales. The first one, originated by Mr. George E. Cooper, is practically neglected today. It is based on a 1 to 10 rating, the number increasing with the deficiencies, and is given in Figure 1. This first scale often led to confusion, depending on the configuration being "Normal" or "Emergency."

A second scale, called the "CAL Rating Scale," did not include this factor. It is given in Figure 2. Although bringing a significant simplification to the write up, its application still leads to some confusion in interpreting the descriptors, and in the fact that it may rate absolutely differently the same test configuration depending on whether the situation is "Normal" or "Emergency." For example, an engine failure may be rated 5 because it is a failure, or 2 because its effects are negligible.

To best eliminate the differences in interpretation and render the rating scale universal without any risk of misinterpretation of the descriptors, Messrs. George Cooper and Robert Harper have devised an excellent rating scale which has been partially taken up by the Franco-British officials in the SST Standard No. 5. This scale is based on a sequential choice process, successively deleting the options until the final choice is made.

The sequential choice is shown in Figure 3. This scale is certainly the surest amongst those in current use, while leaving some gradation liberty in the final choice. It is given in Figure 4.

3. CRITICISM OF THE COOPER-HARPER SCALE

If it is properly used, the Cooper-Harper rating scale should provide a low scatter in the pilots' opinion. It is its application which is difficult.

In fact, once the configuration and the phase of flight are defined, the pilot must make the selection between the normal operation cases and those with failures, and judge for himself, in advance, what he is to accept or not of the proposed configuration before judging the configuration itself.

Thus, for example, when asked "Is adequate performance attainable with a tolerable pilot workload?", he will have to make his opinion on the probability of configuration occurrence, on the cases when it can happen, on its effects relative to continuation of flight, possibly on "average" pilot reactions, etc. This is, of course, rather constraining and therefore rather difficult to obtain, especially when there are many cases to be tested and when the pilots have not been trained to that kind of mental activity.

Here we reach the second objection - the rating sequential process is relatively slow and misadapted to simulator work where the cases to be dealt with are often numerous and in rapid succession. In those circumstances, the pilots "forget" the sequence and switch directly to the final rating number, often with rather personal interpretations.

The third objection is that this scale groups quantatively and unequivocally such concepts as difficulty, workload and safety. These are certainly interrelated, but not so rigidly as required by the Cooper-Harper rating. This has frequently given us some concern in data analysis.

4. CHOICE OF A NEW SCALE

The new scale tries to meet the objections to the Cooper-Harper scale while retaining a qualitative aspect allowing easy data analysis. Amongst other things, it transfers from the pilot to the data processing team a major part of judgment interpretation. Of course, the configuration and the phase of flight must be as well defined as for the Cooper-Harper rating scale, but the rest does not need to be very accurately defined.

Concepts are called for which offer the pilot no ambiguity in interpretation, with very simple ratings going from 1 to 3 in the sense of increasing difficulty. The three basic concepts - skill, attention required and safety - are sufficiently segregated in his mind so that he can rate them separately, even if they are interrelated in the particular situation he is requested to judge.

To allow particularly the pilot to properly judge the required attention, it will, of course, be stated if the flight condition is normal or with failure, and the failure probability can even possibly be mentioned.

Only the use of this rating scale will allow us to tell if its use is efficient in reaching the objectives.

5. H.A.S. RATINGS

Difficulty and Workload Rating	I.	Skill required *(Habileté)*	H	low	1
				average	2
				too important	3
	II.	Attention required *(Attention)*	A	low	1
				average	2
				too important	3
Safety Rating	III.	Safety *(Sécurité)*	S	without risk	1
				"potential" risk	2
				immediate risk	3

6. USE OF H.A.S. RATINGS

Due respect being paid to the precautions mentioned in Section 4, the pilots will rate the flight conditions tested by means of three marks.

It is the analysis of these three marks which should allow one to subsequently judge the conditions and configurations tested by segregating difficulty and safety. As already mentioned, the three marks are not absolutely independent. Thus, a case such as $H = 3$ and $A = 3$ will probably lead to $S = 2$ or $S = 3$.

Inversely, the case $H = 1$ and $A = 1$ will lead to $S = 1$. But one can imagine situations where this is not true - for example, in the case of an approach with application of autothrottle failure. We can have $H = 1$ and $A = 1$ or 2, but $S = 2$, the failure being liable to place the aircraft in a hazardous angle of attack situation if it is not detected early enough, or if the failure occurs at an unfavorable time. The ratings possible with the Cooper-Harper scale for this example could be either 2 or 8.

Discussion of results using the H.A.S. scale could be made on the following basis:

- acceptable cases ($H \leq 2$: $A \leq 2$: $S = 1$)
- questionable cases ($H \leq 3$: $A \leq 3$: $S \leq 2$)
- unacceptable cases ($H = 3$: $A = 3$ with $S \leq 3$)
  or ($S = 3$)
  or ($S = 2$)

depending on the probability of occurrence of the condition.

	Adjective rating	Numerical rating	Description	Primary mission accomplished	Can be landed
NORMAL OPERATION	Satisfactory	1	Excellent, includes optimum	Yes	Yes
		2	Good, pleasant to fly	Yes	Yes
		3	Satisfactory, but with some mildly unpleasant characteristics	Yes	Yes
EMERGENCY OPERATION	Unsatisfactory	4	Acceptable, but with unpleasant characteristics	Yes	Yes
		5	Unacceptable for normal operation	Doubtful	Yes
		6	Acceptable for emergency condition only	Doubtful	Yes
NO OPERATION	Unacceptable	7	Unacceptable even for emergency condition	No	Doubtful
		8	Unacceptable - Dangerous	No	No
		9	Unacceptable - Uncontrollable	No	No
	Unprintable	10	"Motions possibly violent enough to prevent pilot escape"		

Figure 1. Original Cooper Rating Scale

Category	Adjective description within category	Numerical rating
Acceptable and satisfactory	Excellent Good Fair	1 2 3
Acceptable but unsatisfactory	Fair Poor Bad	4 5 6
Unacceptable	Bad Very bad Dangerous	7 8 9
Unflyable	Unflyable	10

Figure 2. CAL Rating Scale

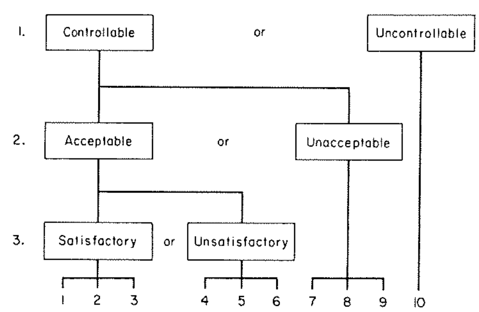

Figure 3. Sequential Pilot-Rating Decisions

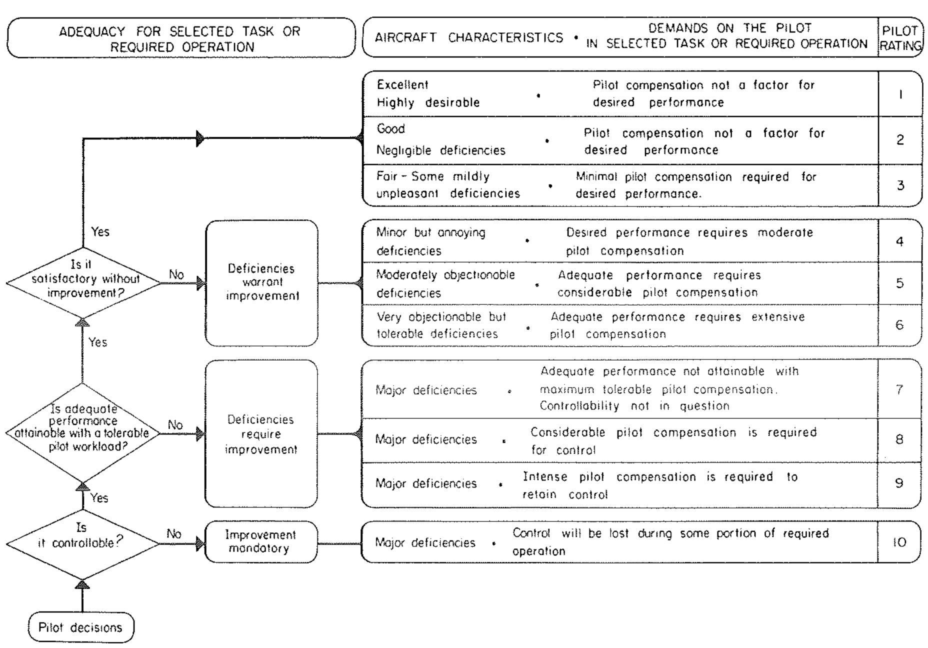

Figure 4. Handling Qualities Rating Scale

LEAD DISCUSSION

by

Lt. C. A. Wheal, RN
Royal Aircraft Establishment, U.K.

I arrived at this Symposium with a prepared lead discussion and a lot of preconceived ideas. In the course of the Symposium the preconceived ideas have gone out the window and this would be the third time I have rewritten the lead discussion if, in fact, I had had time to do more than sketch out a few thoughts on paper. I find it encouraging that the last word at this Symposium has been given to the pilots, because it makes the point that all the expense, effort and brainpower going into producing better simulators is aimed ultimately at producing better aeroplanes or making better use of the ones we have. The man who decides the success or failure of this aim is, in the long run, the pilot: a fact which will remain true until all flight profiles are automated and the simulation engineer's dream comes true - the pilot is completely described by a mathematical function. There are so many points raised in the paper by Messrs. Cooper and Drinkwater that it would be impossible to discuss more than a very few of them in detail, so I intend to pick one or two points which I feel particularly strongly about and perhaps look at them from a slightly different point of view.

1. EXTRAPOLATION

The first of these points is the question of extrapolation, a theme which runs through the whole paper prepared by Messrs. Cooper and Drinkwater. Indeed, in a sense it has been the theme of the whole Symposium, since in its most general meaning extrapolation is concerned with the extension of simulator results to real flight. If that could not be done there would be no justification for simulators at all. Here, however, I should like to concentrate on the aspect that Messrs. Cooper and Drinkwater mentioned, namely pilot extrapolation, which by implication is confined to research applications. I must point out that my remarks will not necessarily apply to such advanced research simulators as the Boeing SST project which Mr. Lee described in his paper earlier in this Symposium, as I have never flown anything comparable. Nonetheless I believe my remarks are generally true.

From the point of view of handling research simulation, the value of the simulation lies in trying to stay a little ahead of the flight test programme so that extensions of the flight envelope are not undertaken blindly. The trouble is that the areas of interest are always, by definition, beyond the scope of the validation already performed and the updating from flight results which is constantly taking place. Between simulation and reality lies a gap consisting of all the simplifications, linearizations, unpredictable effects and straight-forward unknowns which are inherent in any simulation of this sort. In addition, the real danger areas are the very places where non-linearities and unpredictable effects are most likely to occur. The question then arises - since the simulator is never really the same as the aircraft, is it better or worse, and by how much? In a marginal case where the simulator is barely controllable, are, for example, the particular motion cues provided in the simulation making a fair situation worse, or an uncontrollable situation manageable? In practice it seems true that in most cases the simulator gives a pessimistic answer and allowances are made for the probability that things will be better in flight. I would very much like to hear Mr. Pinet's comments on the relevance of this to the Concorde programme.

So it seems that in the areas where extrapolation from the simulator to real life would be most valuable, the simulator is least reliable. All the pilot can legitimately be asked to do is to assess the simulator. In a situation where even the scientists involved in a programme admit they do not know how closely the aircraft will resemble the simulation, it is surely unfair to place the pilot in a position of making speculative judgements.

The question then arises: if pilot extrapolation is illegitimate and any other form is dubious, what can be done with research simulation? The value of research simulations is well proved and their usefulness will go on increasing as aerodynamic knowledge increases. At present they provide good general guides as to the sort of behaviour which can be expected from a new aircraft, and clearly as the technology of the subject improves and computations become more flexible and more accurate, the guides will, hopefully, become more precise. One important corollary which has arisen on several occasions is that if a handling problem occurs in flight as predicted in the simulator, there is a good chance that a solution which is successful in the simulator will also be successful in the air.

2. COMPENSATION

Extrapolation, in the paper by Messrs. Cooper and Drinkwater, covers by implication something which I would prefer to call compensation, and this raises a question which I think is of fundamental importance in evaluating the success of any simulation. For the purpose of discussion I should like broadly to define compensation as the process of adjustment by the pilot to simulator deficiencies such as imperfections in motion and visual systems, cockpit interface features, and so on - deficiencies which are not inherent in the aircraft dynamics. The process of compensation may include the development of piloting techniques or the use of cues which differ from those employed in flight. As Mr. Staples pointed out in his paper, human beings are very good at supplying missing information provided the signposts are powerful enough (perhaps we could call this "positive compensation"), but are very bad at suppressing misleading or incorrect information, which we could call "negative compensation". Examples are the familiar optical illusions which Mr. Staples showed, and the quite lengthy flying training process of teaching pilots to ignore misleading vestibular cues when flying on instruments. If we are going to provide a basic or rudimentary simulation in which we rely on the pilot to fill in the gaps in the illusion we are creating by suggestion, it is important that the signposts we use are all pointing in the right direction. It is better to provide no cues at all than misleading ones. Note that this idea of compensation is specifically different from the learning process by which a pilot teaches himself to cope with handling deficiencies in the vehicle being simulated.

If you accept this rather vague concept, it follows that there is a limit of acceptable compensation beyond which the inability to provide missing information or the inability to suppress misleading information adversely affects the pilot's task performance or his ability to make objective evaluations. Many apparently unrelated problems can then be resolved in terms of the level of compensation required before the pilot can perform the task adequately or evaluate handling qualities objectively. For example, simulation fidelity could be determined by the extent to which compensation interfered with the evaluation of the task performance. The adequacy of the simulation for the programme's aims could be determined by similar methods: here, perhaps, one could say that when positive compensation becomes extrapolation the results should become suspect. Excessive negative compensation, of course, would result in the simulation being ruled out for poor fidelity.

The cockpit interface is another area in which excessive compensation is often demanded. The flight instrument arrangement may well bear no resemblance to the proposed aircraft layout, but this is immaterial as long as the instruments are easily read, sensibly laid out and sufficiently sensitive in the important ranges. If, on the other hand, successful task performance requires a high rate of data assimilation in a critical phase of flight - for example, a decelerating transition to a spot vertical landing - then poorly placed or hard-to-read instruments are considerably likely to degrade the pilot's performance. In a case like this the pilot may never by able to give a true assessment of the difficulty of the task or the aircraft's handling qualities because he is unable to imagine (or extrapolate) what it might be like without the built-in handicap of the instruments. The only question which I hope nobody asks is, how do you recognise the limit of acceptable compensation? All I can say is "Ask the pilot", and the more pilots you ask the more answers you will get.

3. LANDING PERFORMANCE

Many speakers have mentioned the degraded landing performance achieved on simulators. Clearly this is ultimately attributable to shortcomings in the simulation, predominantly in the visual and motion cues, but it occurs to me that the relationship between cause and effect may not be as simple as it seems at first sight. There seems to be a tacit assumption that the pilot is trying just as hard to achieve good landing performance in the simulator as he would in an actual aircraft. If this were true, the degradation of performance would be directly related in some way to the simulator's deficiencies. But is it true? Landing an aircraft manually is a closed-loop process in which the pilot is operating at very high gain. A recent study by Pinsker suggests that this is a condition of potential instability which is only alleviated by the fact that the pilot's position is ahead of the aircraft's center-of-gravity, providing him with phase-advanced, high-gain visual feedback cues. It is arguable that a similar analysis could be applied to normal acceleration cues. In any case it is clear that any reduction of gain or introduction of unwanted lags in the feedback loops will increase the tendency towards instability. Instability of this sort, resulting in a pilot-induced oscillation, is uncomfortable and bad for the pilot's morale, so he reduces his own gain, making the whole process tend towards an open-loop manoeuvre. The performance degradation is still a measure of the simulator's deficiencies but the relationship is complicated by the extent to which the pilot is willing to approach his own instability boundary.

4. WORKLOAD

I was going to say quite a lot about workload since I have recently been involved in a programme of measuring pilot heart rates during jet V/STOL aircraft flight, but the subject has been hammered to death at this Symposium and all the things I wanted to say have either been said or disproved. However, there is one aspect of workload which has not received the attention it deserves, and that is the provision of a realistic R/T task in training simulators. There are some notable exceptions - for example, the hookup of SST simulations with real ATC networks, though this was done for different reasons - but in general nobody seems to have considered it. As any pilot will testify, keeping a listening watch through poor reception, clipped transmissions and chatter from other aircraft can make up a significant portion of the total workload. Could I therefore make a plea for some consideration to be given to this problem in connection with training simulators?

5. SIMULATOR VALIDATION

The problems of simulator validation are quite different in the fields of training and research simulators. A research simulator may be flown exclusively or predominantly by test pilots, who are not noted for their reticence when confronted by a simulation which does not fly like an aeroplane. In this case a good deal of the allotted programme time may have to be devoted to getting the simulation right, and once the engineers have established that the general sense of operation of components is correct, the cost-effective answer is to enlist the test pilot's help at the earliest opportunity.

When the simulation is to be flown by non-test pilots, it is even more important that it should be validated by test pilots before the programme starts. If this is not done, the assessment pilots may be called upon to compensate for deficiencies which could affect the results of the experiment. This may never be discovered, as non-test pilots have not been trained to observe and analyse what they see, and are likely to be reluctant to criticise a simulation about which they probably know little.

With training simulators the problems are twofold - the validation must be as accurate as possible, and in addition, when the simulator is to be used for routine proficiency checks, crew rating and so on, the authorizing body may insist that certain performance criteria are met, usually at the corners of the flight envelope. These requirements may be incompatible - because of the need to simulate systems performance, full mission capability, navigation or nav-attack facilities and so on, the limitations of computer capacity may impose simplifications which do not permit the aircraft's behaviour to be simulated accurately throughout the envelope. I have heard of simulators in which one set of derivatives was used to meet the certification criteria and another set sold to the customer. There is a need for much wider education of the people concerned as to just what it is reasonable to expect of a simulator.

The other part of the question concerns the overall fidelity of the simulation, which must obviously be as high as possible. At present, most of the validation and acceptance trails of training simulators are carried out by pilots with no special qualifications for the task other than experience on type or seniority in the organization, or both. No one doubts the good intentions on the part of the pilots or the firms, but the acceptance pilots may not be able, because they have not been trained, to isolate and analyse faults or to discuss them in terms the engineers can understand. As Mr. Breuhaus has said, when the pilot says something is wrong you'd better listen - but how do you listen to somebody who does not speak the same language? As a result, many training simulators enter service with a number of more or less serious faults, most of which could be completely and quickly eliminated by letting an experienced simulator test pilot perform the validation.

This policy has two major consequences. The first is that every training simulator which enters or remains in service in a state of poor or incomplete validation is a permanently bad advertisement which will, because of the large number of pilots who use it, add considerably to the already widespread prejudice against simulators on the part of aircrews. The second is that most pilots outside the research world have limited simulator experience. If they have never flown a really good one, they have no idea of the potential of even the current generation of simulators. Consequently, acceptance standards are lower than they need be and the manufacturers, who are not in business for altruistic reasons, are under no pressure to develop or improve their products. I except the increasing size, cost and complexity which are always on offer to the customer whose requirements are principally for greater prestige - apart from being more expensive, a bad six-degree-of-freedom motion system is likely to be about twice as bad as a bad three-degree-of-freedom system, and the provision of such a system does not automatically constitute either development or improvement.

6. VISUAL DISPLAYS

The point has been made several times during this Symposium that the development of visual displays in the last few years has not kept pace with that of motion systems. I had a vivid demonstration of this yesterday afternoon, when thanks to Mr. Bray and Mr. Dusterberry, I was privileged to fly the FSAA here at NASA Ames. This is a very expensive simulator with a very advanced motion system which, although not completely optimised yet, provides high fidelity cues of a quality outside my previous experience. The collimated colour projection visual system, on the other hand, is no advance on the systems which have been used in countless other simulations for years. It is as good as the best of its kind, and it is certainly not intended as adverse criticism to say that it is nonetheless the weak link of this simulator. The fact is, we just do not know how to build a better visual display.

It is not my intention to do more than voice a few thoughts on the subject, but there are some aspects which deserve mention. George Cooper has asked whether colour is essential in visual systems. When it is correctly adjusted, a good colour display is very good, but it does have its own problems such as the difficulty of obtaining and keeping colour registration and focus, the apparently poor depth of field, the increased maintenance time, and so on. A badly adjusted colour system is worse than a black and white display - few scenes look less realistic than one where the lights have haloes, the grass is purple, and the sky is red. Colour visual displays are like multi-axis motion systems - the potential is there but is unlikely to be realized unless great care is taken in their setting up and maintenance.

The future of visual displays is obscure. For all the experimentation with point-light sources, shadowgraphs, multi-tube TV displays, anamorphic film projections, and so on, it is clear that in the present state-of-the-art it is nothing more than experimentation - no system has significant overall advantages and certainly none provides the complete answer. Personally I share with Doug McGregor the hope that eventually laser technology will enable a completely realistic holographic display to be developed.

A final point. Amidst all the talk at this Symposium about redundant, distorted or missing visual cues and the admissions that nobody knows just how a pilot looks at the visual scene and interprets what he sees, let alone how he uses what he interprets, has anyone thought about night deck landings? On a black and horizonless night there is a complete absence of peripheral cues, streamer cues or textural cues - in fact, the whole external visual scene consists of the deck landing sight, centre line lights and a horizontal reference on the ramp. The information obtainable from such a limited scene is obviously sufficient, but equally obviously there is no redundancy involved. This real life visual scene bears a close resemblance to some contact-analogue displays, and prompts one to ask the question that if we suspect a TV-model display of subtly and imperceptibly distorting the visual cues used by the pilot and thereby degrading his performance, might we not, with advantage, consider putting greater effort into the development of contact-analogue systems?

7. TURBULENCE

Finally, a few remarks on turbulence. The representation of turbulence in simulators often comes in for some harsh criticism from pilots, in most cases quite justifiably. The trouble lies mainly not with the short or the long wavelength disturbances which produce respectively predominantly structural excitation and flight path disturbances, but with the medium wavelength disturbances which produce predominantly rotational aircraft responses. The simulator's response to this type of turbulence is often exaggerated badly enough to reduce significantly the realism. The reason is not hard to find - for perfectly valid reasons most simulators use a point mass, rigid body representation of the aircraft for turbulence response. This ignores the fact of progressive aircraft immersion in an atmospheric disturbance and produces an exaggerated incidence response. It also precludes asymmetric lateral immersion and the resultant rolling response, unless this is fed in separately. It may be possible to improve this aspect of simulation by applying a scale factor to all disturbance signals of the appropriate wavelengths. In any event, it should not be ignored because like most of the things we have been discussing, a good turbulence representation may not be singled out for especial notice, but a bad one is a continual reminder to the pilot that he is in an unreal situation - in other words, that the aircraft he is flying is "only a simulator".

OPEN DISCUSSION

Cooper (USA)

Mr. Pinet's comments are very interesting because they are based on simulation assessments that are closely related to a flight development program, and the simulation grew with the project. Many of the remarks in our paper relate to general experiments on handling qualities. I compliment Mr. Pinet on his new approach to a rating scale. I feel you must get more and more specific in the questions you ask the pilot, and this is a limitation with any rating scale. On the third classification in Mr. Pinet's rating scale - safety - I wonder how effective a pilot can be in assessing this particular item?

Hurley (USA)

We must not omit the needs of training simulators from our discussions since there is a growing overlap between training and research simulators. Two years ago the Air Transport Association emphasized the poor quality of the data supplied for previous training simulators and highlighted the need to improve training simulators using research simulator technology. As a consequence of the concern expressed by the ATA, a considerable effort by the airframe manufacturers has gone into training simulators. The Boeing 747 training simulators are the first to reflect this effort, and all but one of these have six-degree-of-freedom motion systems. We have taken considerable care with regard to the data on performance - including aerodynamics and propulsion - as well as stability and control characteristics, the flight control system and display instrumentation. Boeing and other companies have spent from 2 1/2 to 3 million dollars just to generate the data needed for the training simulators. For the Boeing 747 there is only one summary aerodynamic report and it is contained in the training simulator document. The same is true for the summary report on the engines, which are modelled very completely with extensive nonlinearities. We believe that it is only in the simulator that we can get together the complete description of the aircraft in one place.

Doetsch (Germany)

Has the requirement for improved training simulator fidelity increased the amount of data the aerodynamicists and aircraft systems engineers had to generate?

Hurley (USA)

Yes, it has, and it also has required more flight testing. On the Boeing 747 we flew quite a few flight test hours just to get the data needed for the training simulator. We photographed the display instruments during engine starts, in-flight shut-downs, and other situations and plotted the readings in order to get the correct dynamics in the simulator. We also ran special wind tunnel tests on such things as asymmetric flap deflections and partial landing gear extensions and incorporated these characteristics in the computer. In addition, we went back and flew about a half million dollars worth of flight tests on the Boeing 707 to obtain better data for the 707 simulator.

Burny (Belgium)

We must have good data not only for the research simulators but also for the training simulators. I agree with Mr. Hurley on the importance of accurate data for training simulators. On the subject of workload, it seems to me that we must try to arrive at the point where the workload in the simulator will be the same as in the aircraft.

Cooper (USA)

I agree that in the final analysis we would want the same workload in the simulator. But this is a difficult aspect. In the real situation the workload fluctuates, and it is essential to be able to assess the most critical situation. Referring to the time for distraction which I mentioned previously and which I feel is the key parameter, perhaps you might not have the same workload in the simulator but could have varying times of distraction - that is, distract the pilot 25%, 50% or 75% of the time until his performance begins to break down.

Beyer (Germany)

I would like to ask Mr. Pinet it he intends to expand his rating scale to multi-dimensions. Also, is Mr. Cooper doing more work on rating scales?

Pinet (France)

We divided the scale into three parts because from our experience we found that it was difficult to determine what the pilot really meant when he gave just one rating.

Sometimes the single rating was based on task performance and sometimes it was based on what dangers could arise. The pilots were influenced by different considerations in giving their rating. We find it better to rate separately the three factors. Our scale is multi-dimensional now, since there are situations where the skill and attention levels are low but the aircraft safety may not be too good - for example, an autothrottle approach. In fact, "Skill" could be regarded as a pilot's output and "Attention" as his input.

Cooper (USA)

At one time I seriously considered introducing a skill term in the rating scale, but it would have to be included throughout the scale and not separately for only one or two ratings as originally proposed. This lengthens and complicates the rating descriptions. The problem I had was related to not knowing how to assess the relative levels of skill for different pilots. Also, with regard to Mr. Pinet's scale, we had considered including safety in the rating descriptions, but for simplicity, it was left out as we felt that pilots considered it inherent in the separate ratings as described.

Pinet (France)

Let me define what we mean by the different safety levels. For safety, a rating of 1 means there is no danger at all. Rating 2 means the pilot thinks there is a potential danger - for example, in an autothrottle approach if the pilot is distracted the aircraft may get into a condition where recovery by the pilot could be very difficult. A rating of 3 means there is immediate danger - this rating is close to a 10 on the Cooper scale.